Three-Dimensional Structure of the Second Cysteine-Rich Repeat from the Human Low-Density Lipoprotein Receptor[†,‡]

Norelle L. Daly, Julianne T. Djordjevic, Paulus A. Kroon, and Ross Smith*

Centre for Protein Structure, Function and Engineering, Biochemistry Department, The University of Queensland, QLD 4072, Australia

Received July 11, 1995; Revised Manuscript Received August 28, 1995[⊗]

ABSTRACT: The ligand-binding domain of the low-density lipoprotein receptor comprises seven cysteine-rich repeats, which have been highly conserved through evolution. This domain mediates interactions of the receptor with two lipoprotein apoproteins, apo E and apo B-100, putatively through a calcium-dependent association of the ligands with a cluster of acidic residues on the receptor. The second repeat (rLB2) of the receptor binding domain has been expressed as a thrombin-cleavable GST fusion protein, cleaved, and purified. On oxidation the protein refolded to give a single peak on reverse-phase HPLC. The aqueous solution structure of rLB2 has been determined using two-dimensional ^{1}H NMR spectroscopy. In contrast to the amino-terminal repeat, rLB1, rLB2 has a very flexible structure in water. However, the conformation of rLB2 is markedly more ordered in the presence of a 4-fold molar excess of calcium chloride; the proton resonance dispersion and the number of NOESY cross-peaks are greatly enhanced. The three-dimensional structure of rLB2, obtained from the NMR data by molecular geometry and restrained molecular dynamics methods, parallels that of rLB1, with an amino-terminal hairpin structure followed by a succession of turns. However, there are clear differences in the backbone topology and structural flexibility. As for rLB1, the acidic residues are clustered on one face of the module. The side chain of Asp 37, which is part of a completely conserved SDE sequence thought to be involved in ligand binding, is buried, as is its counterpart (Asp 36) in rLB1. These results provide the first experimental support for the hypothesis that each of the repeats in the ligand-binding domain has a similar global fold but also highlight significant differences in structure and internal dynamics.

The low-density lipoprotein (LDL)[1] receptor is a transmembrane glycoprotein which is involved in cholesterol homeostasis (Goldstein et al., 1979). Apolipoprotein (apo) B-100-containing LDL and apo E-containing lipoproteins, including very low-density lipoprotein (VLDL), intermediate density lipoprotein, and β-VLDL, are internalized by the LDL receptor, thus removing cholesterol and other lipids from the circulation (Mahley & Innerarity, 1983; Brown & Goldstein, 1983). Defects in this receptor result in familial hypercholesterolemia (FH), a common cause of high blood cholesterol levels, which can lead to accelerated atherosclerosis (Brown & Goldstein, 1986).

The receptor is a mosaic protein composed of five discrete functional domains, including a ligand-binding domain containing seven copies of a cysteine-rich repeat (Yamamoto et al., 1984). A diverse range of proteins contain this repeat, including the VLDL receptor (Sakai et al., 1994), the LDL receptor-related protein/α_2-macroglobulin receptor (Herz et al., 1988), renal glycoprotein gp 330 (Raychowdhury et al., 1989), the C8α (Rao et al., 1987) and C9 components of complement (Discipio et al., 1984), the linker protein of earthworm hemoglobin (Suzuki & Riggs, 1993), and a receptor for subgroup-A Rous sarcoma virus (Bates et al., 1993).

The conserved residues throughout these repeats include a cluster of C-terminal acidic residues and the six cysteine residues (Yamamoto et al., 1984). The spacings between the third, fourth, and fifth cysteines and two of the hydrophobic residues, a phenylalanine on the N-terminal half and an isoleucine located in the middle of the repeats, are also highly conserved. Within the repeats of the LDL receptor the C-terminal acidic residues have been postulated to interact with lysine and arginine residues present in the ligands apo B-100 and apo E (Goldstein et al., 1985; Knott et al., 1985; Yamamoto et al., 1984). In addition, calcium binding sites in the repeats have been implicated, as ligand binding is calcium-dependent (van Driel et al., 1987).

Analysis of the ligand-binding domain has revealed that the repeats are not functionally equivalent and that binding of apo E and apo B-100 requires different combinations of repeats. Binding to LDL is affected by deletion of any single repeat, excluding that at the N-terminus (Esser et al., 1988). Deletion of single repeats does not appear to alter the binding to β-VLDL, with the exception of repeat 5, the deletion of which reduces β-VLDL binding by 60%. A naturally occurring mutation which involves the deletion of both the first and second repeats results in normal binding to β-VLDL; however, binding to LDL is reduced by 29% (Esser et al., 1988).

[†] This work was supported by a grant to P.A.K. and R.S. from the Australian National Health and Medical Research Council.

[‡] The atomic coordinates have been deposited in the Brookhaven Protein Data Bank and assigned the identity code 1LDR.

* Address for correspondence: Ross Smith, Biochemistry Department, University of Queensland, QLD 4072, Australia. Tel: 61 7 3365 4627. FAX: 61 7 3365 4699. E-mail: ross@biosci.uq.edu.au.

[⊗] Abstract published in *Advance ACS Abstracts,* October 1, 1995.

[1] Abbreviations: apo, apolipoprotein; CD, circular dichroism; LDL, low-density lipoprotein; rLB1, recombinant N-terminal cysteine-rich repeat of the LDL receptor; rLB2, recombinant second repeat of the LDL receptor; NOE, nuclear Overhauser enhancement; VLDL, very low-density lipoprotein.

0006-2960/95/0434-14474$09.00/0

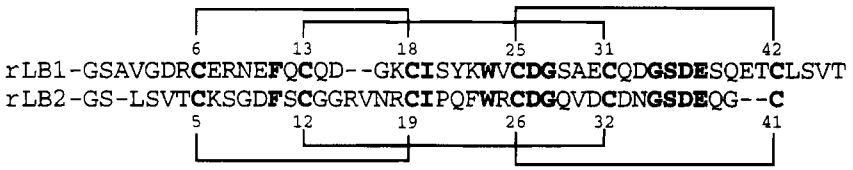

FIGURE 1: The sequences of the recombinant repeats, rLB1 and rLB2. The numbering system used does not include the N-terminal Gly−Ser which is present as a result of the thrombin-cleavage sites in the GST fusion proteins. Conserved residues between the two repeats are in bold type. The cysteine residues are numbered, and the disulfide-bonding connectivities displayed for each peptide.

The three-dimensional structure of the N-terminal repeat of the LDL receptor (rLB1) cleaved from a GST fusion protein has recently been determined by NMR spectroscopy (Daly et al., 1995). Although rLB1 is not involved in binding of either ligand, it was initially chosen for study because of the availability of a monoclonal antibody, IgG-C7, which provides the best available assay for adoption of a native structure by rLB1. The cleaved protein was found to fold to a unique structure that was recognized by the conformationally-sensitive antibody only in the presence of calcium ions. The three-dimensional structure in aqueous solution contains a β-hairpin followed by a series of turns. Within this structure the disulfide bonding was found to link Cys 6 to Cys 18, Cys 13 to Cys 31, and Cys 25 to Cys 42. While no substantial changes occur in the NMR spectra upon addition of calcium chloride to rLB1, this does not preclude the possibility of minor structural changes that are required for antibody binding. Given the sequence homology within the repeats, the expectation was that this structure would serve as a model for the other cysteine-rich repeats.

The second repeat of the LDL receptor (rLB2) has now been expressed as a cleavable GST fusion protein. No conformation-specific antibody is available to examine the folding of rLB2; however, the disulfide-bonding pattern of rLB2, determined by selective reduction experiments, is identical to rLB1 (Bieri et al., 1995a). The sequences and disulfide bonding patterns of rLB1 and rLB2 are given in Figure 1. In the present study we report the structure of rLB2 in aqueous solution.

MATERIALS AND METHODS

The cDNA encoding residues 43−83 of the mature human LDL receptor (rLB2) was PCR-amplified and cloned into the expression vector pGEX-2T (S. Bieri et al., unpublished information). Recombinant LB2 was expressed as a thrombin-cleavable glutathione *S*-transferase fusion protein in *Escherichia coli* using a 15 L Chemap fermentor (Chemap AG, Männedorf, Switzerland). The fusion protein was bound to glutathione−agarose affinity resin chromatography, reduced by incubation with 1 mM DTT for 1 h at 37 °C, and cleaved with bovine thrombin [5 U/(mg of fusion protein)] to release rLB2. Cleavage of the fusion protein with thrombin resulted in an extra two residues (glycine and serine) at the N-terminus of rLB2. The peptide was oxidized in the presence of 3 mM reduced and 0.3 mM oxidized glutathione and subsequently purified by reverse-phase high-performance liquid chromatography (HPLC) using a 8 mm × 100 mm Waters Radial-Pak column (Milford, MA). Gradients of 0.1% aqueous trifluoroacetic acid and 0.1% trifluoroacetic acid in acetonitrile were employed with detection at 214 and 280 nm. Analysis was performed on a Sciex (Thornhill, Ontario) triple-quadrupole mass spectrometer using electrospray sample ionization. The disulfide connectivities were established using selective reduction (S. Bieri et al. 1995a). Quantification of the peptide was achieved using an extinction coefficient at 280 nm of 6050 M^{-1} cm^{-1} (Gill & von Hippel, 1989).

CD spectra were recorded on a Jasco J-710 spectropolarimeter (Tokyo, Japan) over the wavelength range 260−185 nm using a 0.1 mm path length cell, a bandwidth of 1.0 nm, a response time of 2 s, a scan rate of 20 nm min^{-1}, and averaging over 2 scans. The concentration of rLB2 was 220 μM, and the pH adjusted to pH 4.8 with 1 M NaOH. In addition, the effect of Ca^{2+} on the CD spectra was examined by adding calcium chloride to give a final concentration of 870 μM. Spectra were also recorded without calcium in the presence of 30 mM EDTA. For use in the program Varslc1 (Johnson, 1988), a spectrum was recorded from 260 to 178 nm with a 4-s response time and a scan speed of 2 nm min^{-1}. This spectrum was recorded on a sample of rLB2 at 350 μM, pH 5.5, in the presence of an approximately 4-fold molar excess of calcium chloride.

For the NMR experiments rLB2 was dissolved in 10% ^2H_2O, pH 4 or pH 5.5 at a concentration of approximately 1 mM. For experiments in ^2H_2O the sample was lyophilized and dissolved in 99.96% ^2H_2O (Wilmad, Buena, NY). The effect of calcium ions on the NMR spectra was also examined by adding a 4-fold molar excess of $CaCl_2$ (~4 mM). ^1H NMR spectra were recorded on a Bruker AMX 500 spectrometer. Standard pulse sequences were used to obtain DQF COSY (Rance, 1987), TOCSY (Braunschweiler & Ernst, 1983) (mixing times 80 and 100 ms), ECOSY (Griesinger et al., 1986), and NOESY (Jeener et al., 1979) (mixing times 150, 200, 250, and 400 ms) spectra. The water signal was suppressed by low-power irradiation during the relaxation delay (2 s) and during the mixing time of the NOESY experiments. Two-dimensional experiments were generally collected into 4096 data points with 512 t_1 increments of 32−64 scans. Five TOCSY spectra collected with 16 scans and 128 t_1 increments and a relaxation delay of 1 s were interleaved with one-dimensional spectra for the ^2H_2O exchange experiments. Spectra were processed using UXNMR (Bruker software). Generally, both dimensions were multiplied by a sine-bell function shifted by 90° and a polynomial baseline correction applied to selected regions. Spectra were recorded at 303 and 313 K to enable assignment of overlapped resonances. Chemical shifts were referenced to internal 3-(trimethylsilyl)propionic acid-d_4 (Fluka Chemie AG, Switzerland).

A set of 55 structures was generated using 319 nuclear Overhauser enhancement (NOE) distance restraints consisting of 188 intraresidue, 55 sequential, 38 short range (≤5 residues), and 38 long range (>5 residues) NOEs, which were classified as strong, medium, and weak and given upper bounds of 2.7, 3.5, and 5.0 Å, respectively. Peak volumes were measured in an NOESY spectrum recorded with a

mixing time of 150 ms using the EASY program (Eccles et al., 1991) and used to derive distance restraints. Standard values of pseudoatom corrections were added (Wüthrich et al., 1983). In addition, 1.5 Å was added to the upper distance limits for methyl protons. The N-terminal glycine and serine had no distance restraints and were not included in the structure calculations. Torsional restraints were applied to two ϕ angles with bounds of $-120 \pm 30°$ for angles with $^3J_{H^N-H^\alpha}$ coupling constants >9 Hz (residues 19 and 27). An ECOSY experiment was used to estimate $^3J_{H^\alpha-H^\beta}$ coupling constants which were used in conjunction with NOE intensities to derive three χ_1 restraints (residues 9, 24, and 32).

The structures were generated on a Sun SPARCstation5 with *ab initio* simulated annealing using the X-PLOR program (Version 3.1, Molecular Simulations Inc., Burlington, MA; Brunger, 1992), starting with a template and random ϕ and ψ angles. The simulated annealing protocol involved an initial energy minimization in a geometric force field (k_{bond} = 500 kcal mol^{-1} Å^{-2}, k_{ang} and $k_{improper}$ = 500 kcal mol^{-1} rad^{-2}) with a purely repulsive, nonbonded energy term (repel = 1, weight = 0.002), a soft square-well NOE restraining function (k_{NOE} = 50 kcal mol^{-1} Å^{-2}, asymptote = 0.1), and an energy constant of 5 kcal mol^{-1} rad^{-2} for the experimental dihedral angle restraints. A total of 30 ps of high-temperature (1000 K) dynamics was carried out in two stages in which the van der Waals radii, weighting terms, and NOE asymptote were varied. In the first 20 ps of dynamics the weightings used were as follows: van der Waals (0.002), angle (0.4), and improper terms (0.1). These weightings and the NOE asymptote were increased to 1 during the final 10 ps of high-temperature dynamics. Structures were then cooled to 100 K over 15 ps during which time the van der Waals radii and weighting were gradually changed to give final values of 0.8 and 4.0, respectively. The energy constant used for the experimental dihedral angle restraints was increased to 200 kcal mol^{-1} rad^{-2} for this step. Refinement of the structures was achieved with a second round of simulated annealing from 1000 to 100 K over 10 ps using a square-well NOE restraining function (k_{NOE} = 50 kcal mol^{-1} Å^{-2}) and an energy constant of 200 kcal mol^{-1} rad^{-2} for the experimental dihedral angle restraints. The disulfide connectivities (Cys 5−Cys 19, Cys 12−Cys 32, Cys 26−Cys 41) were included as pseudo-NOE restraints during the first round of simulated annealing and explicitly defined during refinement. All peptide bonds were constrained to be *trans*.

RESULTS

Thrombolytic cleavage of GSH−agarose-bound GST-rLB2 resulted in the release of rLB2. The peptide was oxidized at concentrations of ~0.1 g/L in the presence of oxidized and reduced glutathione, yielding fully disulfide-bridged protein that migrated as one predominant peak on HPLC in water/acetonitrile gradients containing 0.1% TFA. The molecular mass of the purified peptide was shown to be 4583.

Circular dichroism spectroscopy was used to establish the conformational stability and the overall secondary structure of rLB2, prior to collection of NMR spectra. The spectrum of rLB2 showed a negative peak near 195 nm, just below the minimum wavelength observed for rLB1 (Figure 2A). As for rLB1, other solution conditions had little influence

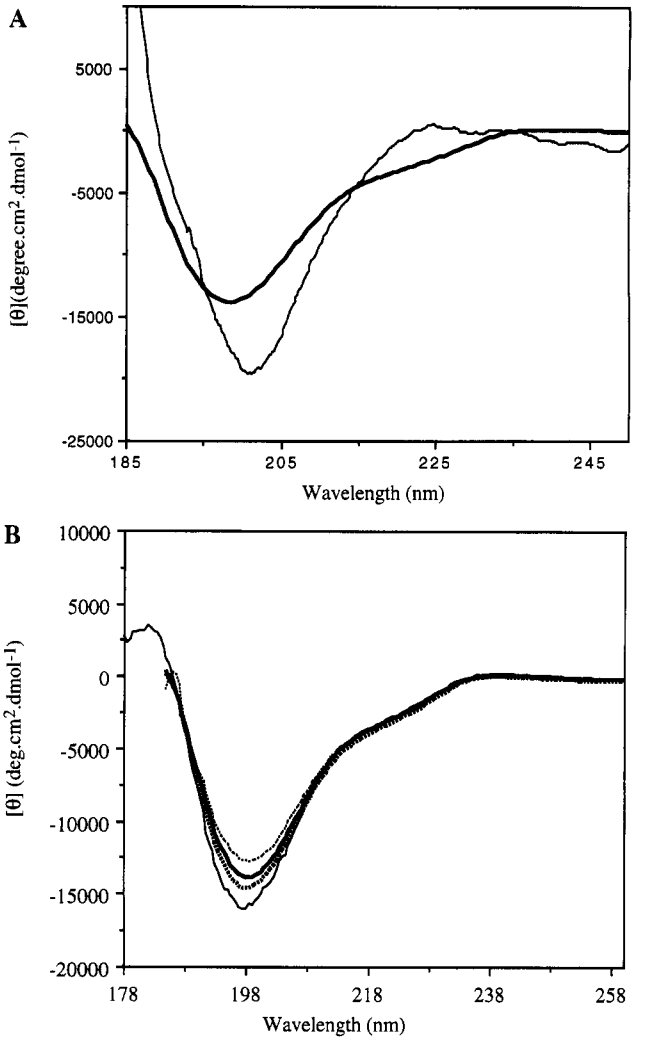

FIGURE 2: (A) Comparison of the CD spectra of rLB1(45 μM, pH 6.9; thin line) and rLB2 (220 μM, pH 4.8; thick line) without calcium present. (B) CD spectra of rLB2, 220 μM at pH 4.8 (thick line), in the presence of 870 μM $CaCl_2$ (thick dashed line), and 30 mM EDTA (dashed line), and at pH 5.5 in the presence of an approximately 4-fold molar excess of calcium chloride (spectrum recorded to 178 nm; thin line).

on the CD spectrum of rLB2: addition or removal of a substantial molar excess of calcium, addition of EDTA, and change in pH from 4.8 to 5.5, by this criterion did not alter the protein structure significantly (Figure 2B). An estimate of the secondary structure was obtained by fitting its CD spectrum to those of basis spectra, using the program Varslc1 (Johnson, 1988) (Figure 2B). This analysis resulted in 13% α-helix, 20% antiparallel β-sheet, 24% turn, and 40% random structure. The overall similarity of the spectra for rLB1 and rLB2 is also in accord with the observed three-dimensional structures.

In contrast to rLB1, the ^{1}H NMR spectra of rLB2 in aqueous solution showed little dispersion, and very few NOE interactions were observed in NOESY spectra with mixing times up to 400 ms. The low dispersion was particularly marked for the amide resonances, which were dispersed over ~2.5 ppm (Figure 3A). There were too few NOESY interactions to allow assignment of the resonances and hence calculation of a three-dimensional structure.

Because calcium is required for the LDL receptor interaction with its ligands, the effect of calcium on the structure

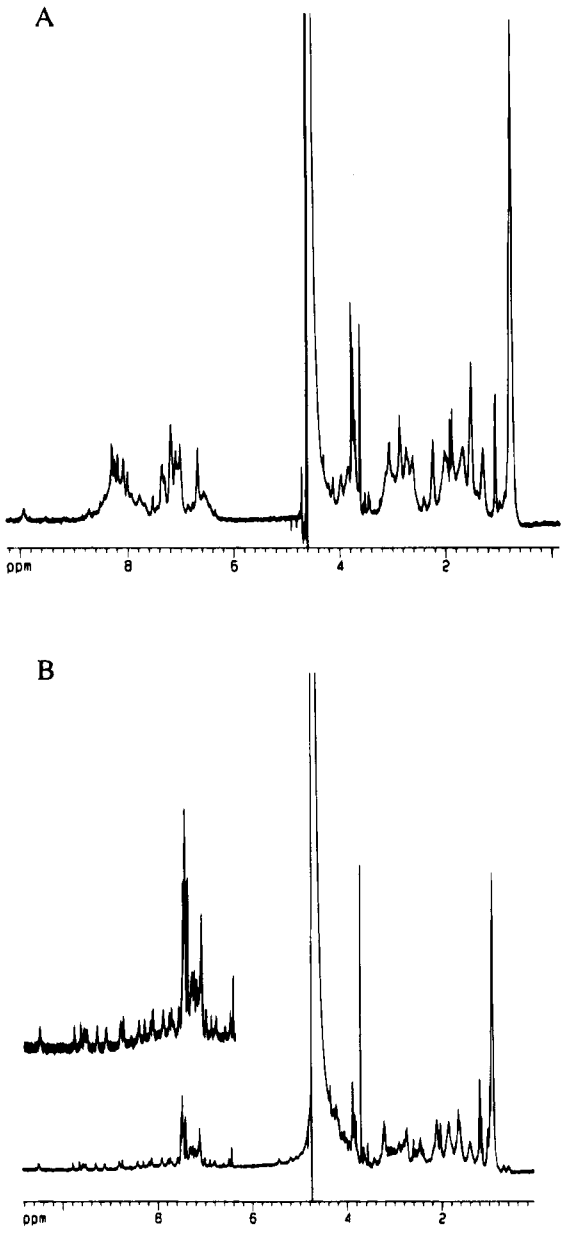

FIGURE 3: One-dimensional ^{1}H spectra of rLB2 (~1 mM, pH 5.5) recorded on a Bruker AMX 500 spectrometer, in the absence (A) and presence (B) of calcium chloride (~4 mM). The amide region intensity is expanded approximately three fold in the insert on spectrum B.

of rLB2 was explored, with effects of unexpected magnitude. As demonstrated in Figure 3B, the dispersion of resonances, and particularly the amide resonances, was greatly enhanced, with several α-proton peaks being shifted to downfield of the water resonance. In addition, a much larger number of NOEs, including medium and long-range interactions, were observed in NOESY spectra recorded with mixing times of 150 ms, enabling sequence-specific assignments and the determination of distance restraints using the method developed by Wüthrich (1986). The effect of calcium on the NMR spectrum of rLB2 is similar to the effect pH had on the spectrum of rLB1. At pH 3.3 (Figure 4A) the amide resonances overlapped, but were well dispersed in NMR spectra of rLB1 recorded over the pH range 3.9−6.5 (Figure 4B).

The short-range NOEs are summarized in Figure 5. A complete set of αN and NN NOES were not observed, not

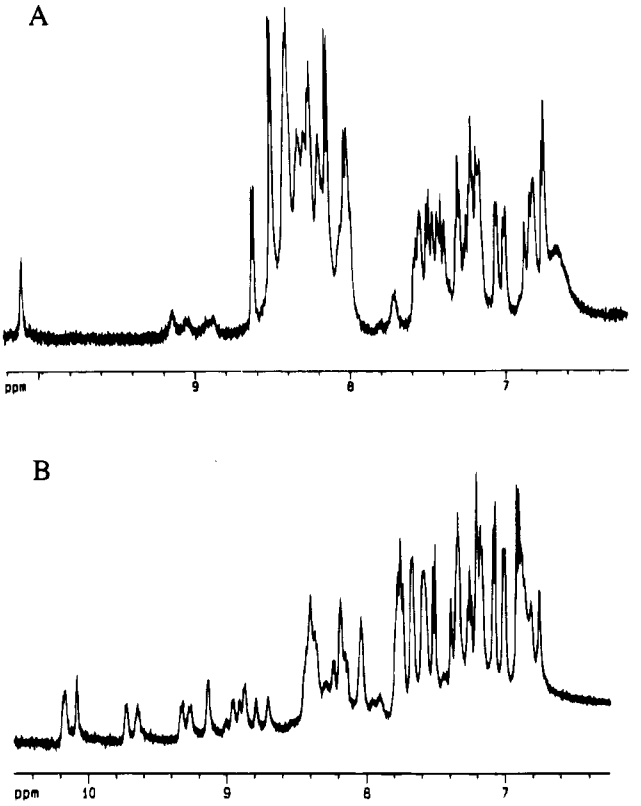

FIGURE 4: Amide region of one-dimensional ^{1}H spectra of rLB1 (4 mM) recorded on a Bruker AMX 500 spectrometer, at pH 3.3 (A) and at pH 6.5 (B). Both spectra were recorded in the absence of calcium chloride.

only because of overlap, but presumably also because of flexibility within the molecule. The intensities of these NOEs varied throughout the molecule, suggesting the presence of irregular secondary structure. The medium-range NOEs between residues 21 and 36 are indicative of turns (Figure 5). The long-range interactions, shown in Figure 6, indicate a folded structure.

From the final set of 55 structures, the 30 structures that had no NOE distance restraint violated by >0.5 Å and no angle restraint violated by >4° were selected for analysis. A superimposition of these structures is displayed in Figure 7. The root-mean-square deviation (RMSD) values from the experimental restraints and idealized geometry were as follows: NOE restraints, 0.07 ± 0.005 Å; dihedral restraints, 0.44 ± 0.55°; bonds, 0.0053 ± 0.0003 Å; angles, 0.74 ± 0.029°; and improper angles, 0.63 ± 0.03°. Energies corresponding to NOE restraints and dihedral restraints were 87.8 ± 13 and 0.15 ± 0.3 kcal mol^{-1}, respectively.

Despite the constraining disulfide bond between residues Cys 5 and Cys 19, the segment amino-terminal to Cys 5 gave no interresidue NOEs and is thus relatively unconstrained in the structures derived from molecular dynamics calculations. Residues 9−12 and 19−32 are well-defined by the NOE restraints and have RMSDs of 0.25 ± 0.21 and 0.42 ± 0.13 Å, respectively, over all backbone heavy atoms. The angle order parameters in these regions, as defined by Hyberts (1992), are close to 1, indicating well-defined angles (Figure 8A). Overall rLB2 is less well defined than rLB1 as shown by the angle order parameters (Figure 8); however, the principal structural features of rLB1 are recognizable in rLB2. A β-hairpin is present in both modules, but with

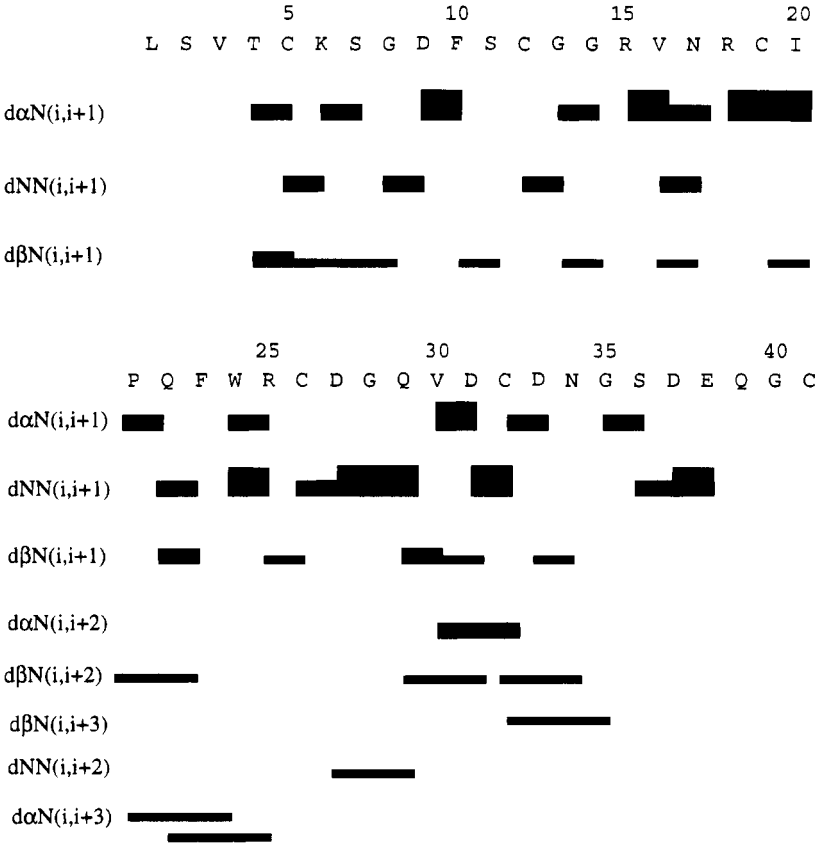

FIGURE 5: Overview of short-range NOEs observed for rLB2. The intensity of NOEs corresponds to the thickness of the lines.

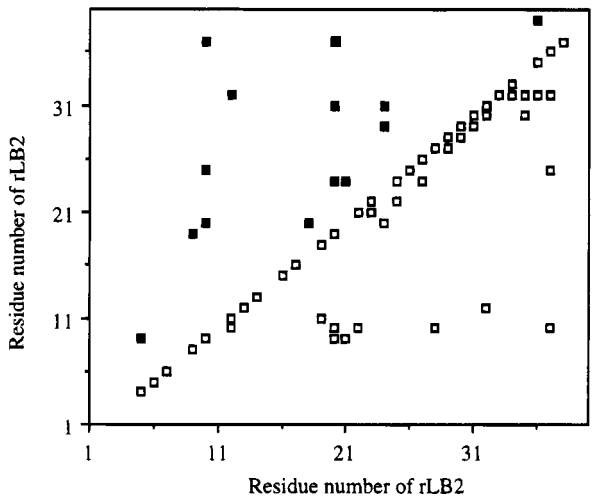

FIGURE 6: Plot of the interresidue NOE distance restraints for rLB2. The open squares represent restraints involving at least one backbone atom, and the filled squares indicate restraints involving only side-chain atoms.

greater flexibility in rLB2, particularly in residues 13−18 which have an RMSD over all backbone heavy atoms of 1.20 ± 0.16 Å. Beyond the hairpin both modules contain a succession of turns, but their detailed topologies and intramolecular motions differ. The structures of rLB1 and rLB2 are compared in Figure 9.

In rLB1 several of the amide protons appeared to be involved in hydrogen bonds, requiring hours for exchange with ^2H_2O (data not shown). By contrast, with rLB2 no amide resonances were present within 10 min of dissolving the protein in ^2H_2O, even in the presence of calcium ions, suggesting that rLB2 is less well constrained than rLB1.

DISCUSSION

For our initial studies of the ligand-binding domain of the LDL receptor we focused on the amino-terminal repeat (Daly et al., 1995; Bieri et al., 1995b). While cognizant of the lack of apparent involvement of this repeat in association with lipoproteins (Esser et al., 1988), we considered it the best starting point because of the availability of an antibody, IgG-C7, that recognizes this repeat within the holoreceptor in a conformationally-dependent manner in the presence of calcium (van Driel et al., 1987; Beisiegel et al., 1981; Daniel et al., 1983). It was reasoned that recognition of the isolated rLB1 module by the antibody in the presence, but not the absence, of calcium ions would constitute evidence for adoption of a similar conformation by rLB1 in the receptor and the isolated module. Folding of rLB1 resulted in one dominant component which was indeed recognized by the antibody. Thus, for rLB1, one arrangement of the disulfide bridges appears to be thermodynamically favored strongly over the alternatives. The same appears to be true for rLB2: although no definitive test for adoption of an *in vivo*-like structure is available, on folding it forms predominantly one conformer, again suggesting that this form is by far the most stable. Significantly, this conformer has an identical disulfide arrangement, and a very similar backbone conformation to rLB1 (Figure 9).

The broad features displayed by rLB1 are retained in rLB2. A flexible amino-terminal segment is followed by a β-hairpin, and then by a succession of turns from residues 22−41. Superimposition of the two backbone structures shows the overall similarity in the fold (Figure 9) but also highlights the differences. The largest single element of secondary structure, the β-hairpin, is present in both, but the additional

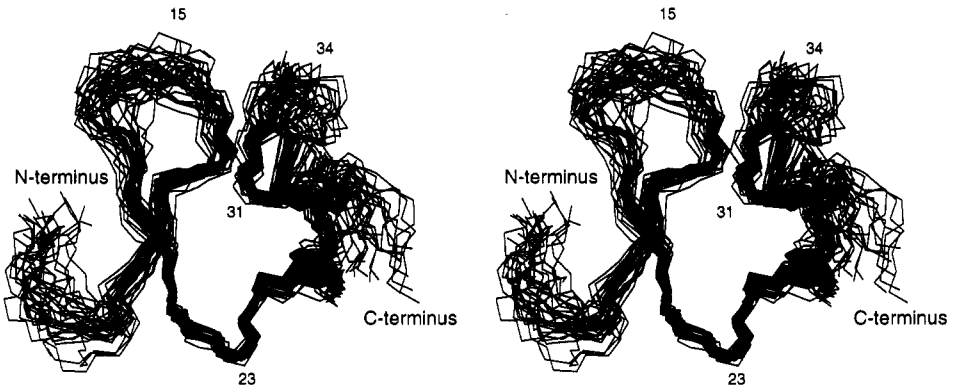

FIGURE 7: Stereoview of the backbone (N, C^{α}, C′) of 30 NMR-derived structures (residues 5−41) of rLB2, superimposed over residues 11−38. The N-terminal Gly−Ser resulting from the inclusion of a thrombin cleavage site in the fusion protein is not included in the structures or in the numbering system.

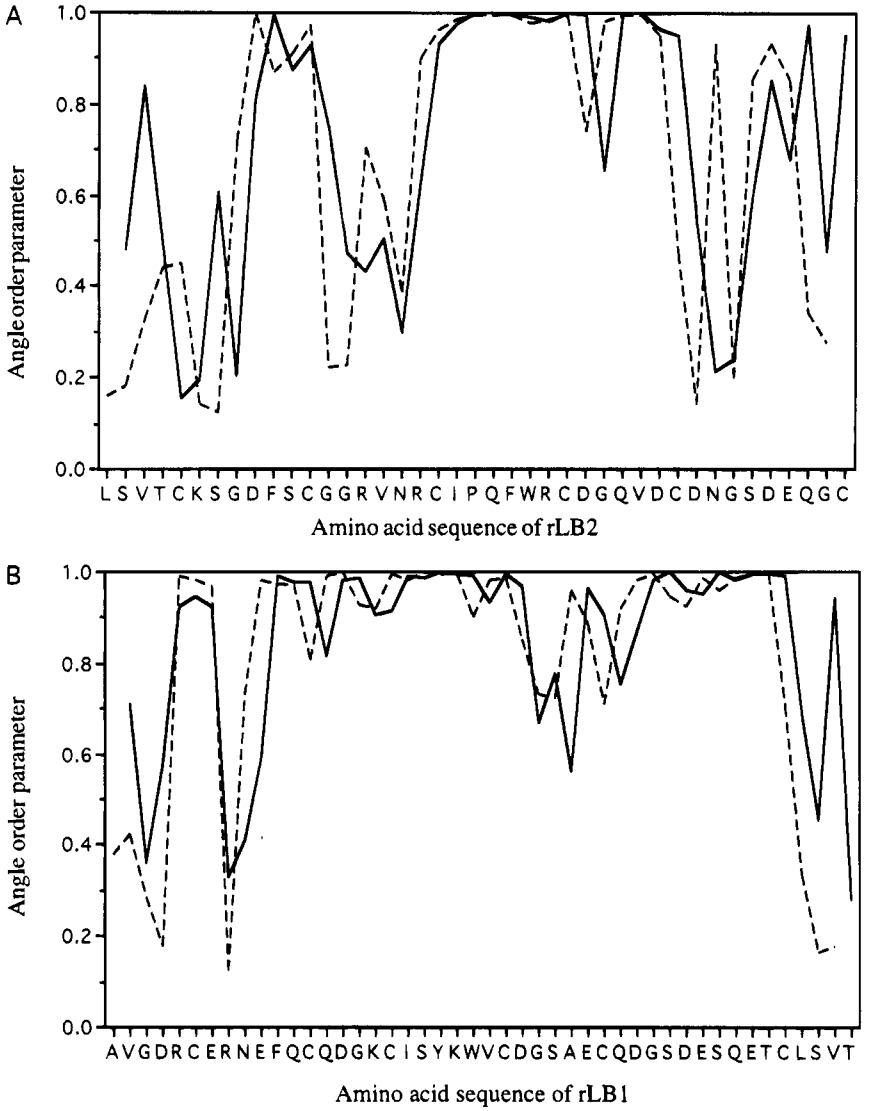

FIGURE 8: The angle order parameters for rLB2 (A) and rLB1 (B), ϕ angle order parameters are shown as solid lines, and φ parameters are shown as dotted lines.

two residues between Cys 12 and Cys 19 in rLB2 cause the hairpin to loop out more, increasing the separation of the two strands with a consequent loss of the interstrand hydrogen bonds.

rLB1 and rLB2 contain proportionately few hydrophobic residues in comparison to most water-soluble, globular proteins, yet rLB1 has a clearly defined hydrophobic core comprising residues Phe 11, Ile 19, Tyr 21, Trp 23, and Val 24. Of the hydrophobic residues in rLB1 only Ala1, Val2, and Leu43 fall outside this core and are exposed to the aqueous solvent: such residues may participate in intermodule interactions in the holoreceptor. Comparison of the se-

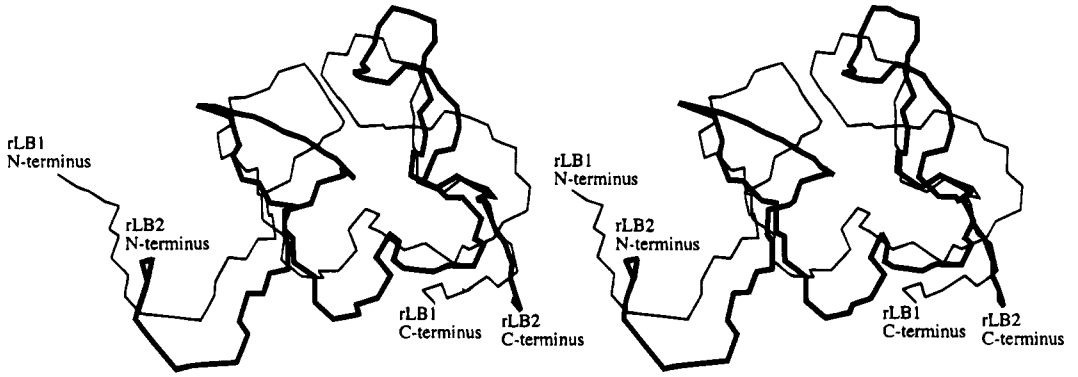

FIGURE 9: Stereoview of a superimposition over residues 10−35 of the backbone atoms (N, C^{α}, C′) of the average rLB1(thin line) and the average rLB2 (thick line) NMR-derived structures (residues 5−41 of both peptides are displayed).

quences of rLB1 and rLB2 shows that rLB2 has residues equivalent in position to residues Phe 11, Ile 19, and Trp 23, but the residue equivalent to Tyr 21 is replaced by Gln and a new hydrophobic residue (Phe) appears in the following position, replacing Lys 22 of rLB1. The Val 24 of rLB1 is replaced by an Arg residue. Of these, the conserved residues Phe (Phe 10 in rLB2), Ile (20), and Trp (24) still form a small core in which the indole side chain remains partly exposed. The shifted Phe 23 is highly solvent exposed, in contradistinction to Tyr 21 of rLB1 which is a part of the hydrophobic core. The volume of this core is further reduced by the loss of a hydrophobic residue following the tryptophan in rLB2. In rLB2 several other hydrophobic residues (Leu 1, Val 3, Val 16, and Val 30) are found outside the core; Leu 1 and Val 30 replace other hydrophobic residues in rLB1, but the other two replace polar or missing residues. These four residues are in contact with solvent, and one can speculate that they may be involved in maintaining the cluster of cysteine-rich repeats in a definite conformation.

The effects of the reduced core of nonpolar residues are evident in the conformational stability of the module. Although constrained by three disulfide bridges, rLB2 is clearly a very flexible structure. In the absence of calcium this intramolecular motion is manifested by a very small dispersion of NMR resonances. In the presence of calcium the three-dimensional structure is far better defined, as judged by both the greatly increased amide resonance dispersion and greater number of NOEs. Yet even in the presence of calcium there are indications that the structure is less constrained than rLB1: whereas rLB1 has several protons that exchange over hours, the protons of rLB2 exchange within the shortest period that could be used in recording the NMR spectra, about 10 min. Thus, the hydrogen bonds which must stabilize the hairpin in rLB1 are absent from rLB2. It is of interest that, despite the dramatic effect of calcium on the latter repeat, the CD spectrum was not changed by addition of this ion, indicating that there is no induction of major new elements of secondary structure. At present the stoichiometry of calcium binding to rLB2, and the identity of the groups to which it is bound, is unknown, but given its influence on the rLB2 structure, it is possible that it enhances the association with ligands by dictating the conformation of the receptor binding domain.

One of the major recent advances in our understanding of the mechanism of protein folding has been the realization that proteins may pass through a stable intermediate on passage from the fully unfolded state to the folded state. In this intermediate, which has been dubbed the "molten globule" state, the protein has most of the secondary structural elements that it has in the fully folded state, but it has a more expanded and flexible tertiary structure. rLB2 in the absence of calcium appears to fit this description as the CD spectra indicate that its secondary structure does not differ greatly from that adopted in the presence of calcium, yet the NMR spectra clearly indicate that with calcium bound rLB2 is markedly different in structure or dynamics. There is a close parallel with the behavior of other calcium-binding proteins such as caltractin (Weber et al., 1994), though these proteins generally have EF hands as their Ca^{2+}-binding motifs, a structure that is not present in the rLB repeats. A calcium-induced transition from a molten globule to the fully folded state has also been proposed for α-lactalbumin (Pardon et al., 1995).

Given the putative role of the acidic residues in ligand binding, it is interesting to look at the disposition of these residues in rLB2, and to compare it with rLB1. The latter has the residues Asp 4, Glu 7, Glu 10, Asp 15, Asp 26, Glu 30, Asp 33, Asp 36, Glu 37, and Glu 40. With the exception of Asp 4 and Glu 7, these acidic residues lie on one face of the module. The same general disposition of acidic residues is evident in rLB2: Asp 27, Asp 31, Asp 37, and Glu 38 again are grouped on one face, but their relative positions do not completely mirror those in rLB1. Intriguingly, the side chain of Asp 36, which is part of an exceptionally conserved SDE cluster that has a putative role in ligand binding, is buried in rLB1 as is its counterpart, Asp 37, in rLB2. In rLB2 this residue may form a salt bridge with Arg 25, but in rLB1 there is no basic residue in close proximity.

In summary, the information encoded within the sequences of rLB1 and rLB2 appears sufficient to direct their folding overwhelmingly to one of the 15 potential arrangements of disulfides. Although folded independently *in vitro*, both reach a cognate three-dimensional structure which must now be considered very likely to be mimicked by the remaining five repeats of the ligand-binding domain.

ACKNOWLEDGMENT

We would like to thank Mr. S. Bieri (University of Queensland) for information on the disulfide connectivities prior to publication and Dr. D. Doak (University of Oxford) for the use of his X-PLOR script for determining the angle order parameters.

SUPPORTING INFORMATION AVAILABLE

One table giving the chemical shifts of rLB2 at 303 K and pH 5.5 (2 pages). Ordering information is given on any current masthead page.

REFERENCES

Bates, P., Young, J. A. T., & Varmus, H. E. (1993) *Cell 74*, 1043–1051.

Beisiegel, U., Schneider, W. J., Goldstein, J. L., Anderson, R. G. W., & Brown, M. S. (1981) *J. Biol. Chem. 256*, 11923–11931.

Bieri, S., Djordjevic, J. T., Jamshidi, N., Smith, R., & Kroon, P. A. (1995a) *FEBS Lett. 371*, 341–344.

Bieri, S., Djordjevic, J. T., Daly, N. L., Smith, R., & Kroon, P. A. (1995b) *Biochemistry 34*, 13059–13065.

Braunschweiler, L., & Ernst, R. R. (1983) *J. Magn. Reson. 53*, 521–528.

Brown, M. S., & Goldstein, J. L. (1983) *J. Clin. Invest. 72*, 743–747.

Brown, M. S., & Goldstein, J. L. (1986) *Science 232*, 34–47.

Brunger, A. T. (1992) *X-PLOR*, Version 3.1, Yale University, New Haven, CT.

Daly, N. L., Scanlon, M. J., Djordjevic, J. T., Kroon, P. A., & Smith, R. (1995) *Proc. Natl. Acad. Sci. U.S.A. 92*, 6334–6338.

Daniel, T. O., Schneider, W. J., Goldstein, J. L., & Brown, M. S. (1983) *J. Biol. Chem. 258*, 4606–4611.

Discipio, R. G., Gehring, M. R., Podack, E. R., Chen Kan, C., Hugli, T. E., & Gey, G. H. (1984) *Proc. Natl. Acad. Sci. U.S.A. 81*, 7298–7302.

Eccles, C., Güntert, P., Billeter, M., & Wüthrich, K. (1991) *J. Biomol. NMR 1*, 111–130.

Esser, V., Limbird, L. E., Brown, M. S., Goldstein, J. L., & Russell, D. W. (1988) *J. Biol. Chem. 263*, 13282–13290.

Gill, S. C., & von Hippel, P. H. (1989) *Anal. Biochem. 182*, 319–326.

Goldstein, J. L., Anderson, R. G. W., & Brown, M. S. (1979) *Nature 279*, 679–685.

Goldstein, J. L., Brown, M. S., Anderson, R. G. W., Russell, D. W., & Schneider, W. J. (1985) *Annu. Rev. Cell. Biol. 1*, 1–39.

Griesinger, C., Sorensen, O. W., & Ernst, R. R. (1986) *J. Chem. Phys. 85*, 6837–6852.

Herz, J., Hamann, U., Rogne, S., Myklebost, O., Gausepohl, H., & Stanley, K. K. (1988) *EMBO J. 7*, 4119–4127.

Hyberts, S. G., Goldberg, M. S., Havel, T. F., & Wagner, G. (1992) *Protein Sci. 1*, 736–751.

Jeener, J., Meier, B. H., Bachmann, P., & Ernst, R. R. (1979) *J. Chem. Phys. 71*, 4546–4553.

Johnson, W. C. (1988) *Annu. Rev. Biophys. Chem. 17*, 145–166.

Knott, T. J., Rall, J., S. C., Innerarity, T. L., Jacobson, S. F., Urdea, M. S., Levy-Wilson, B., Powell, L. M., Pease, R. J., Eddy, R., Nakai, H., Byers, M., Priestley, L. M., Robertson, E., Rall, L. B., Betsholtz, C., Shows, T. B., Mahley, R. W., & Scott, J. (1985) *Science 230*, 37–43.

Mahley, R. W., & Innerarity, T. L. (1983) *Biochim. Biophys. Acta 737*, 197–222.

Pardon, E., Haezebrouck, P., De Baetselier, A., Hooke, S. D., Fancourt, K. T., Desmet, J., Dobson, C. M., Van Dael, H., & Joniau, M. (1995) *J. Biol. Chem. 270*, 10514–10524.

Rance, M. (1987) *J. Magn. Reson. 74*, 557–564.

Rao, A. G., Howard, O. M., Ng, S. C., Whitehead, A. S., Colten, H. R., & Sodetz, J. M. (1987) *Biochemistry 26*, 3556–64.

Raychowdhury, R., Niles, J. L., McCluskey, R. T., & Smith, J. A. (1989) *Science 244*, 1163–1165.

Sakai, J., Hoshino, A., Takahashi, S., Miura, Y., Ishii, H., Suzuki, H., Kawarabayasi, Y., & Yamamoto, T. (1994) *J. Biol. Chem. 269*, 2173–2182.

Suzuki, T., & Riggs, A. (1993) *J. Biol. Chem. 268*, 13548–13555.

van Driel, I. R., Goldstein, J. L., Südhof, T. C., & Brown, M. S. (1987) *J. Biol. Chem. 262*, 17443–17449.

Weber, C., Lee, V. D., Chazin, W. J., & Huang, B. (1994) *J. Biol. Chem. 269*, 15795–15802.

Wüthrich, K., Billeter, M., & Braun, W. (1983) *J. Mol. Biol. 169*, 949–961.

Wüthrich, K. (1986) in *NMR of Proteins and Nucleic Acids*, Wiley, New York.

Yamamoto, T., Davis, C. G., Brown, M. S., Schneider, W. J., Casey, M. L., Goldstein, J. L., & Russell, D. W. (1984) *Cell 39*, 27–38.

BI9515681